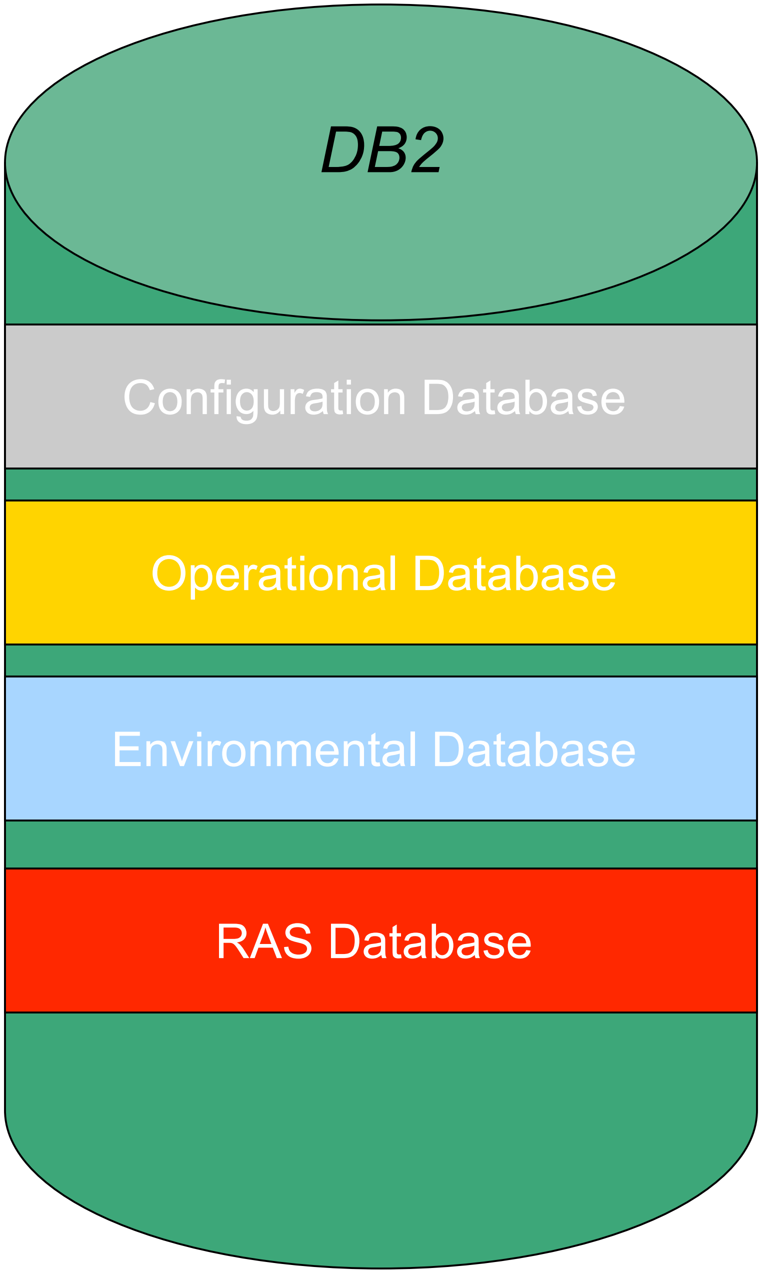
BG/L Control System Software

Mark Megerian
IBM Rochester

February 2005 | Blue Gene/L

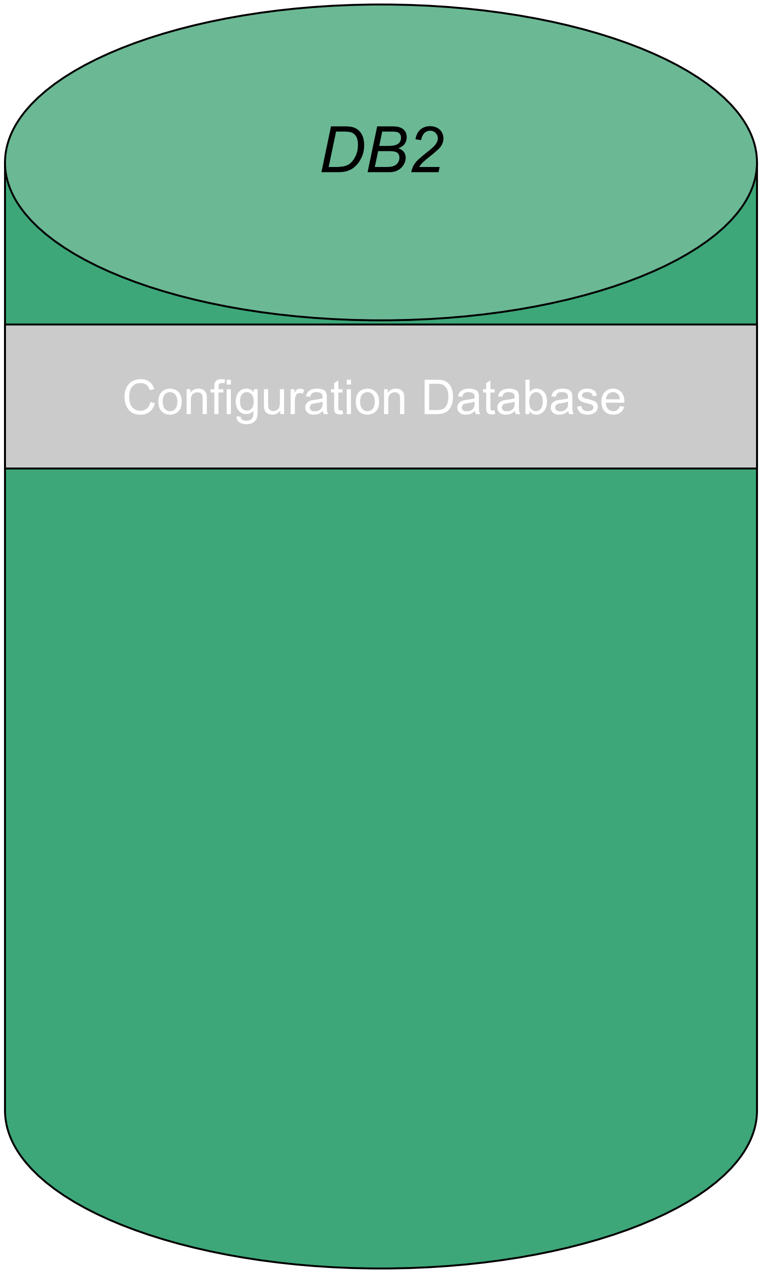
BG/L Control System Overview

- Control system software runs on the service node
- Control system software manages the aspects of system operation up through job submission onto the BG/L core
  - Hardware discovery
  - System partitioning
  - Booting of partitions
  - Job submission / job polling
  - RAS data collection
  - Hardware monitoring
- DB2 is the central repository of all control system information
  - Allows control system components to get hardware information and topology from the database, which is always kept current
  - Less direct contact with the hardware



BG/L DB2 Structure

- Configuration database is the representation of all the hardware on the system
- Operational database contains information and status for things that do not correspond directly to a single piece of hardware such as jobs, partitions, and history
- Environmental database keeps current values for all of hardware components on the system, such as fan speeds, temperatures, voltages
- RAS database collects hard errors, soft errors, machine checks, and software problems detected from the compute complex.



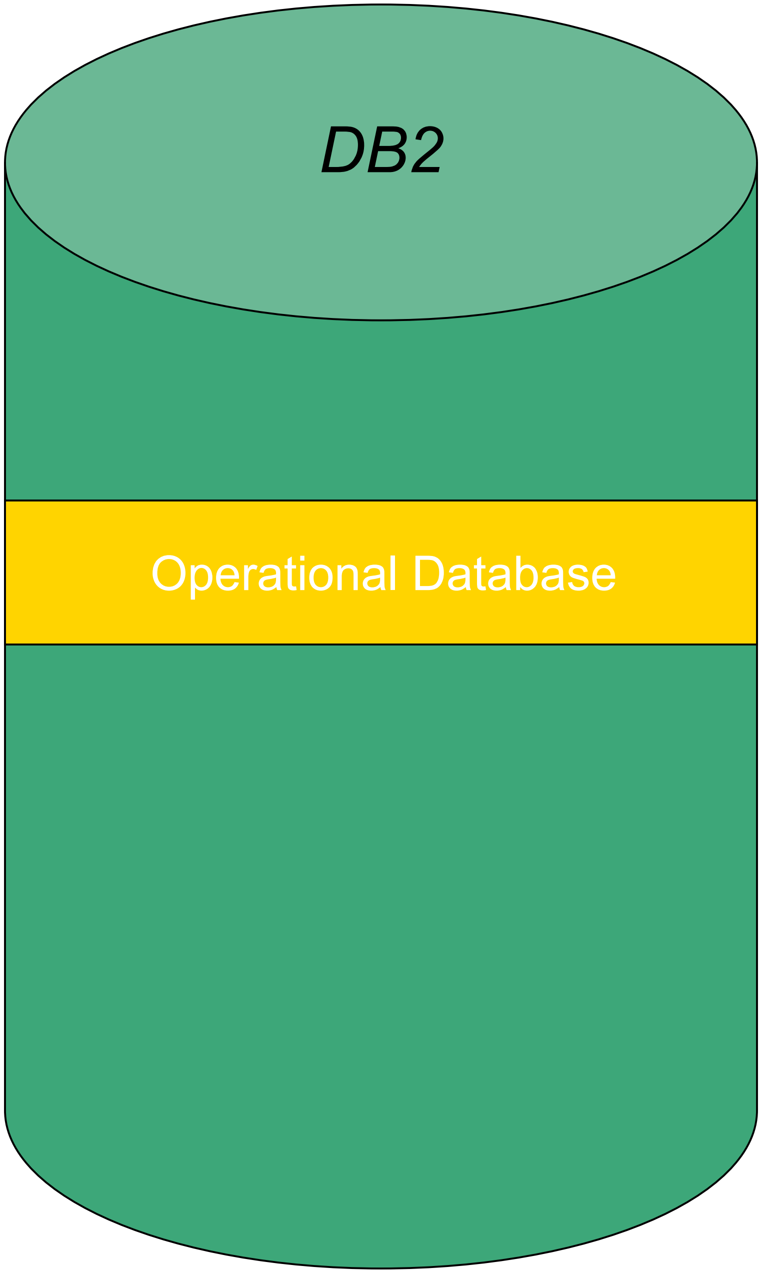
DB2

Configuration Database

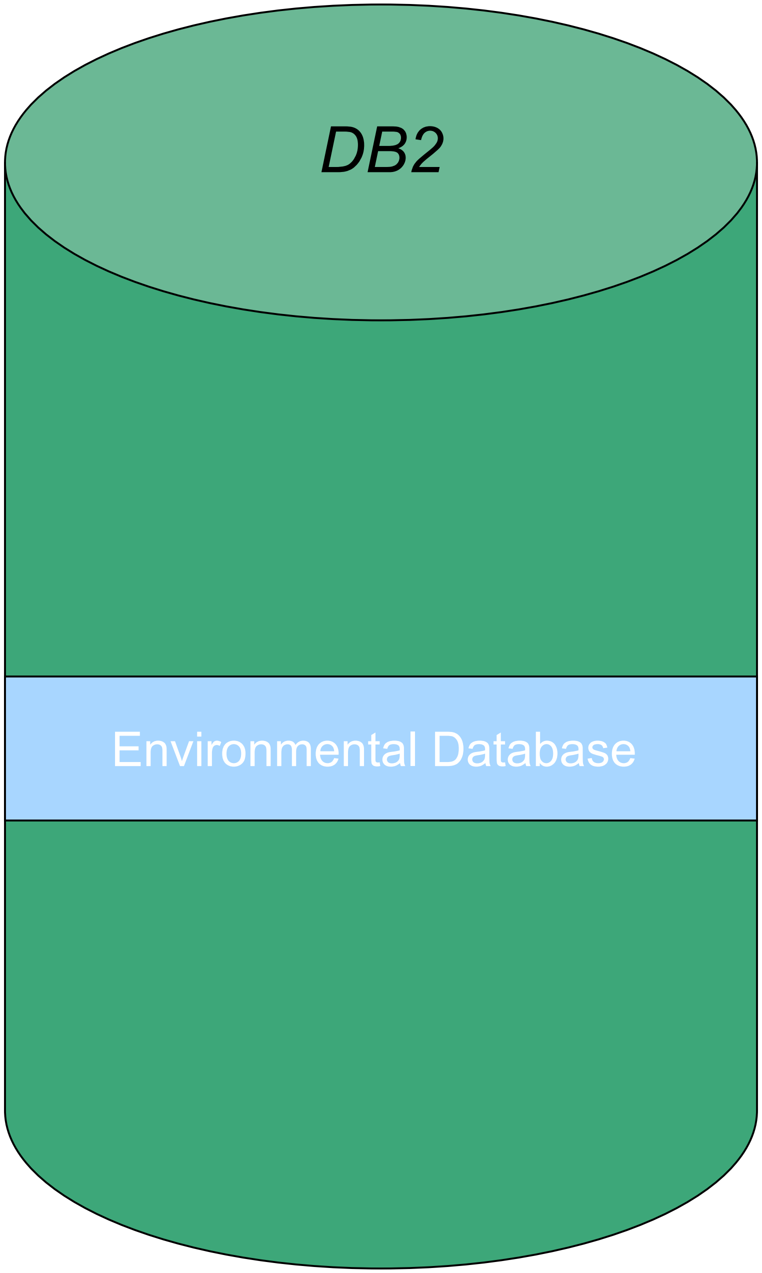
BG/L DB2 Structure

- Configuration database is the representation of all the hardware on the system
  - Machine
  - Midplanes
  - Service Cards
  - Link Cards
  - Link Chips
  - Node Cards
  - Processor Cards
    - Compute & I/O
  - Nodes
  - Cables
  - Ido Chips
  - Clock Cards
  - Fan Modules

BG/L DB2 Structure



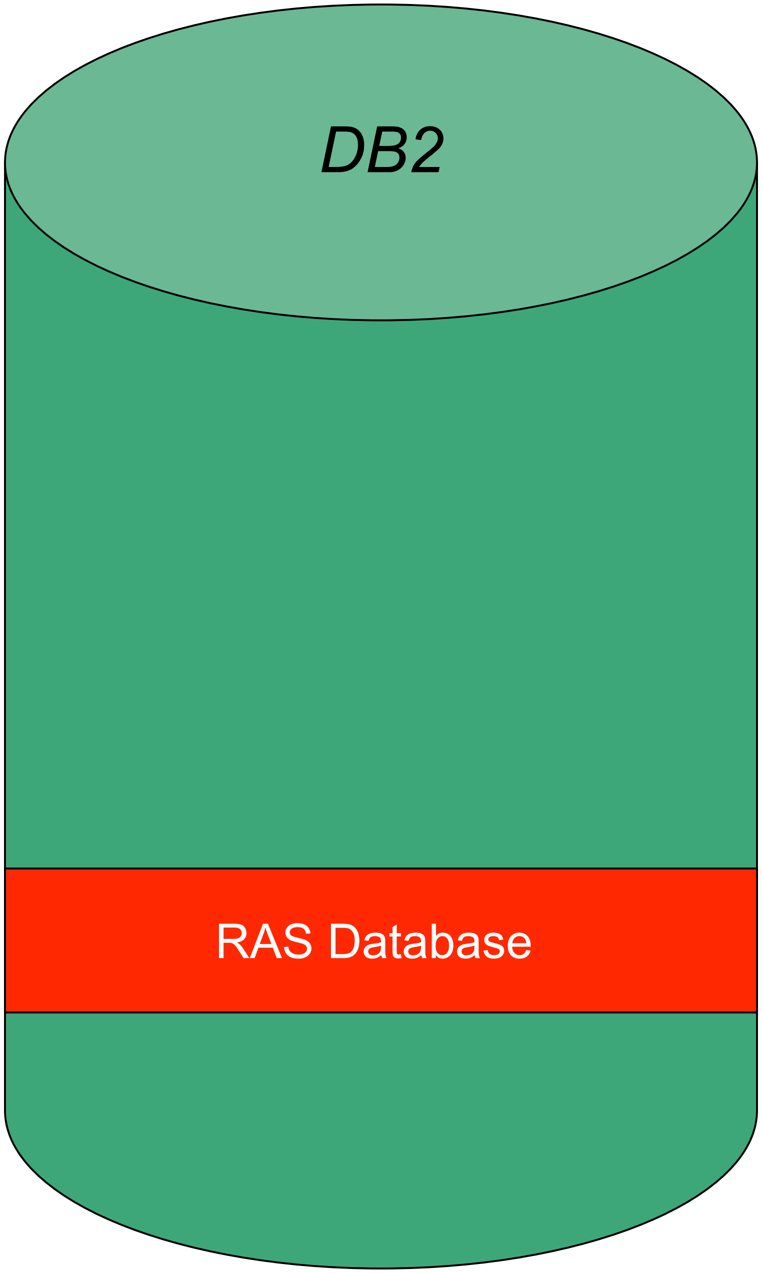
- Operational database contains information and status for things that do not correspond directly to a single piece of hardware such as jobs, partitions, and history
  - Blocks (partitions)
  - Jobs
  - Job history
  - Switch settings
  - Link <-> Block map
  - Block users



BG/L DB2 Structure

- Environmental database keeps current values for all of hardware components on the system, such as fan speeds, temperatures, voltages
    - Fan Modules
        - Desired RPMs
        - Actual RPMs
        - Voltages
        - Temperatures
    - Service Cards
        - Ambient temp
        - Chip temps
        - Voltages
    - Node Cards
        - Chip temps
        - Temp limits
        - Wiring faults
    - Link Cards
        - Power Status
        - Temps

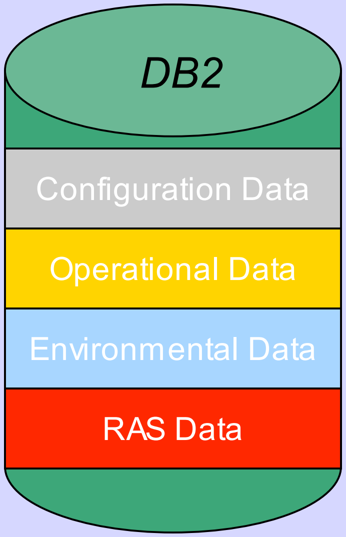
BG/L DB2 Structure



- RAS database collects hard errors, soft errors, machine checks, and software problems detected from the compute complex.
  - RAS events collected by Discovery for bad hardware, missing cards, bad memory, bad cables
  - RAS events collected from compute complex while jobs are running, from kernel interrupts
  - RAS events generated by HW monitoring, for wiring faults, bad cards, fan speeds, over temps
  - RAS events generated by MMCS during link training, software errors, file system errors

Control System Components

BG/L Service Node



DB2

Configuration Data

Operational Data

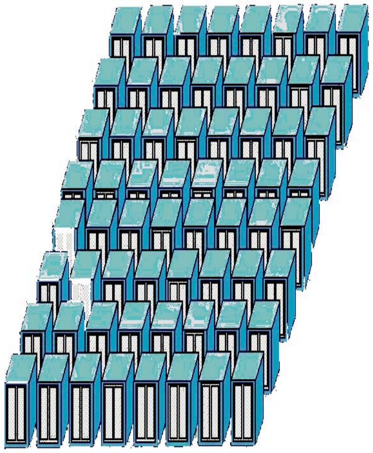
Environmental Data

RAS Data

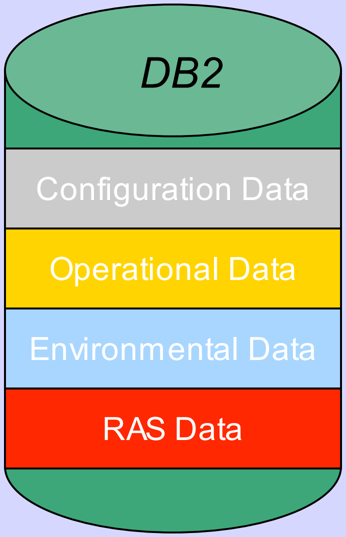
- •Proxy
- •Discovery
- •MMCS
- •RAS
- •Monitor
- •CIODB

All are controlled by a “master” process that watches each process, and restarts upon failure.

All interact directly with the database, and interact with BG/L core.



Control System Components



DB2

Configuration Data

Operational Data

Environmental Data

RAS Data

Proxy

Discovery

MMCS

RAS

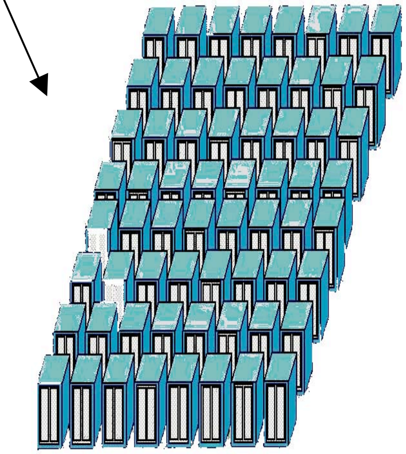
Monitor

CIODB

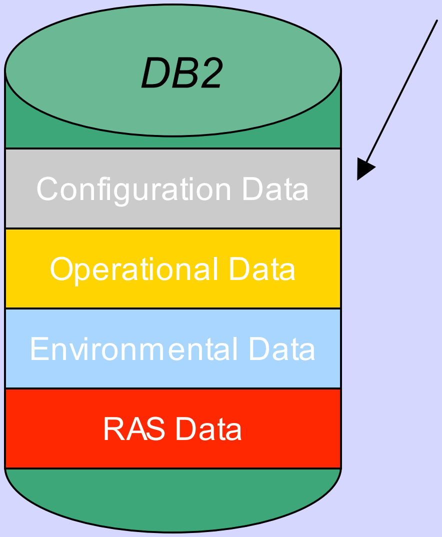
Always running on service node.

Marshals traffic to the hardware, and blocks multiple connections to the same hardware.

Accesses the database for configuration data



Control System Components

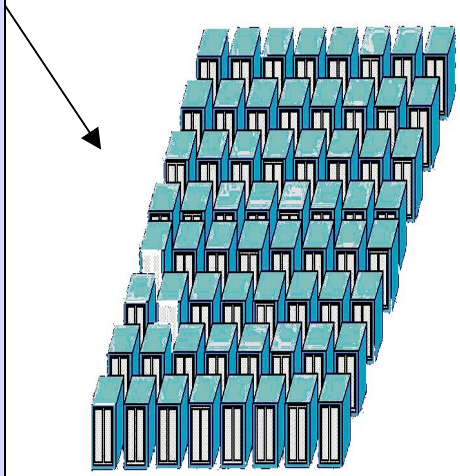


- Proxy
- **Discovery**
- MMCS
- RAS
- Monitor
- CIODB

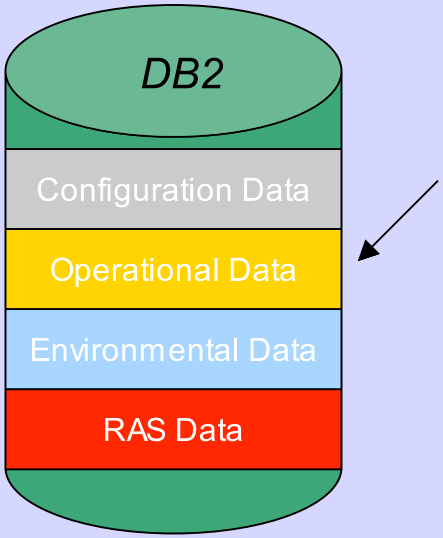
A process that is used during initial setup that discovers all the hardware on the system and populates the database with the configuration info.

Discovers cables between midplanes.

Runs beyond initial setup to flag hardware that goes offline, or new hardware that is added.



Control System Components



Proxy

Discovery

MMCS

RAS

Monitor

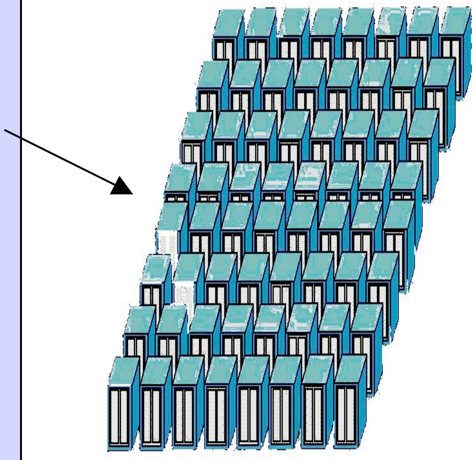
CIODB

Always running, and polling the database.

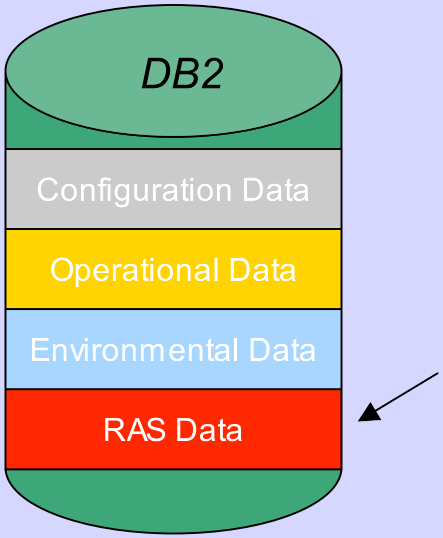
MMCS Server polls for tasks submitted by mpirun, mmcs consoles, and by scripts

MMCS also spawns threads for monitoring each block, once its been booted

There is an MMCS console that can be used to directly interact with MMCS



Control System Components



DB2

Configuration Data

Operational Data

Environmental Data

RAS Data

Proxy

Discovery

MMCS

RAS

Monitor

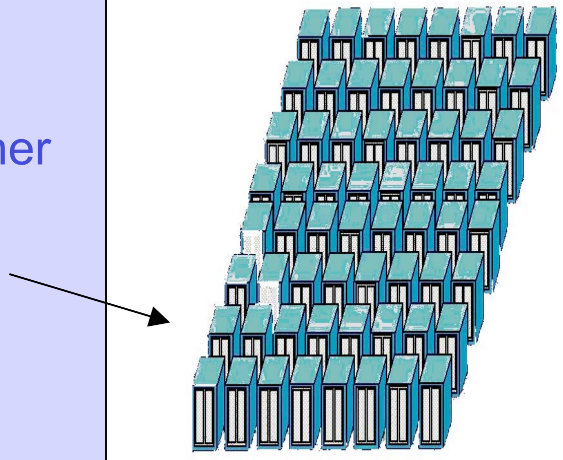
CIODB

Reliability / Availability / Serviceability Data

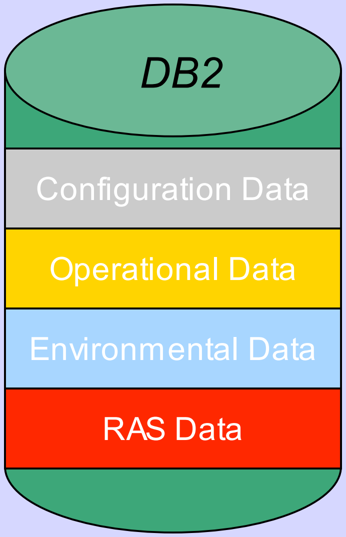
Collected from nodes during job execution

Also collected during other phases, such as booting blocks, discovering hardware, monitoring hardware

All RAS events stored with location of event, timestamp, block, job, node number, event text



Control System Components



Proxy

Discovery

MMCS

RAS

Monitor

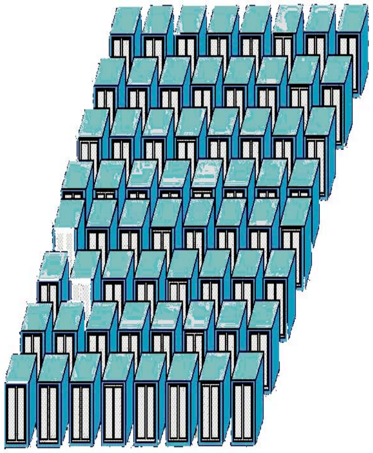
CIODB

Runs all the time, or on an “as-needed” basis.

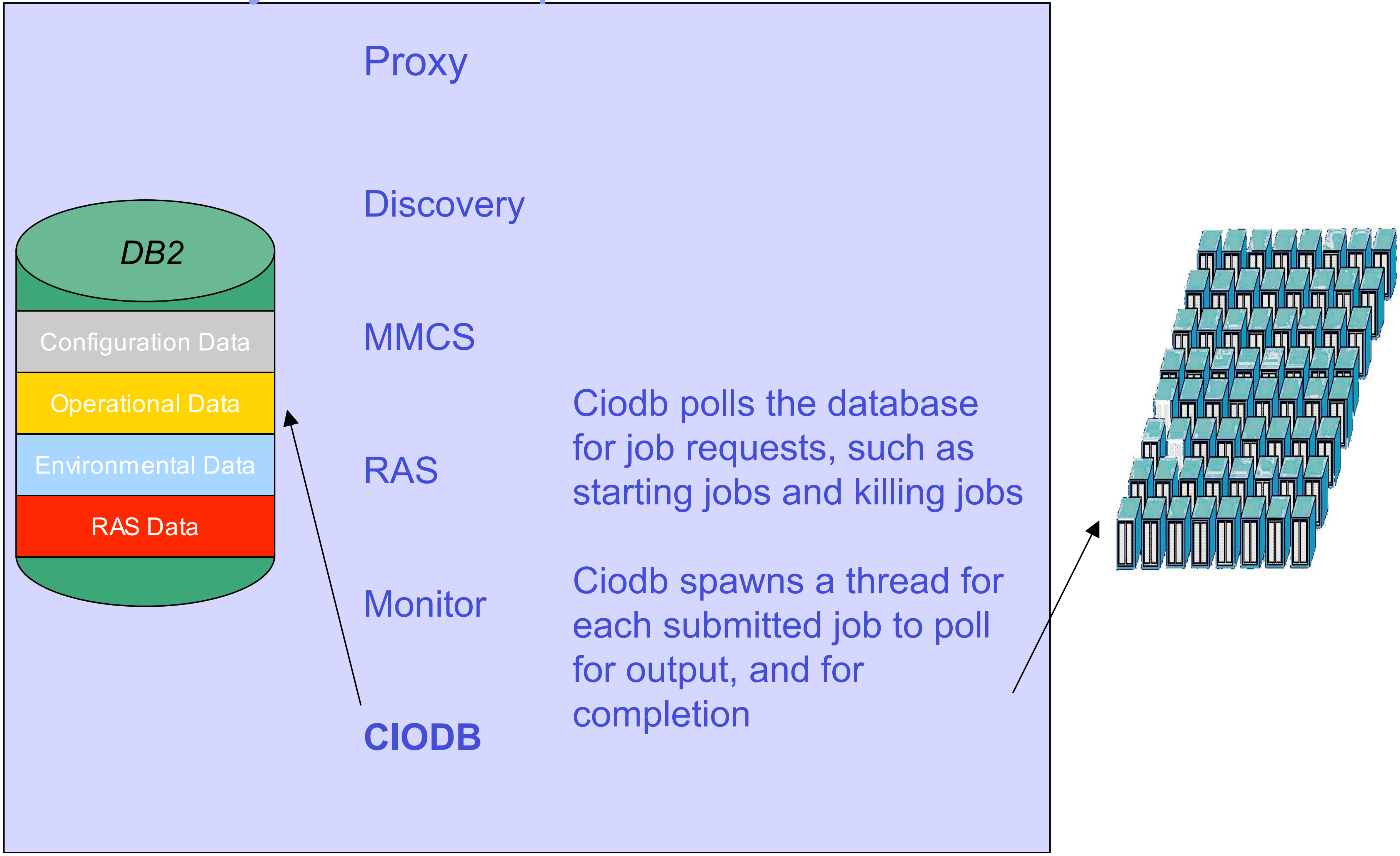
Monitors individual cards for temperatures, fan speeds, voltages, power supply status, etc.

Console for controlling all the monitoring, and viewing the data that has been collected.

Runs against node cards, link cards, and service cards.



Control System Components



DB2

Configuration Data

Operational Data

Environmental Data

RAS Data

Proxy

Discovery

MMCS

RAS

Monitor

CIODB

Ciodb polls the database for job requests, such as starting jobs and killing jobs

Ciodb spawns a thread for each submitted job to poll for output, and for completion

Life Cycle of a Partition

- Partition is defined as a collection of midplanes and switch settings, and implied cables
- Partition that is defined, starts in FREE status, and contains the following information:
  - Block ID (partition name)
  - Size (number of midplanes)
  - Shape (x, y, z dimensions)
  - Torus or Mesh
  - Mode (co-processor mode, virtual node mode…)
  - Ratio of IO nodes to compute nodes
  - Path for Microloader image
  - Path for RAMdisk image
  - Path for Linux kernel image
  - Path for CNK image
  - Creation timestamp
  - Owner
  - Status
- Partitions can be defined using the console, with APIs using XML, by calling MPIRUN, or by using an external scheduler

Life Cycle of a Partition

- Partition is initially FREE
  - The process of booting the partition starts with an “allocate”
- Partition goes to ALLOCATED
  - The components of the partition are allocated, and therefore cannot be used by another partition
  - The ido connections are established for node cards and link cards via the proxy
  - The switch settings are made to program the link chips for either torus or mesh
- Partition goes to CONFIGURING
  - Microloader is loaded onto all nodes
  - RAMdisk is loaded onto IO nodes
  - Linux kernel is loaded onto IO nodes
  - CNK is loaded onto compute nodes
- Partition goes to BOOTING
  - All nodes are started, torus and tree links established
  - IO nodes mount the file system and establish ethernet connections
  - Ciod starts, and sends [ciod initialized] message to CIODB
- Partition goes to INITIALIZED when all IO nodes have responded
  - Jobs can be submitted
  - MMCS polls for RAS events
  - Freeing the partition takes it back to FREE, and ido connections are released

Life Cycle of a Job

- Job is created on the system in QUEUED status
- Job contains the following information
  - Path to executable
  - Partition name
  - User name
  - Arguments
  - Environment variables
  - Stderr and stdout file or redirection
  - Working directory
  - Status

Life Cycle of a Job

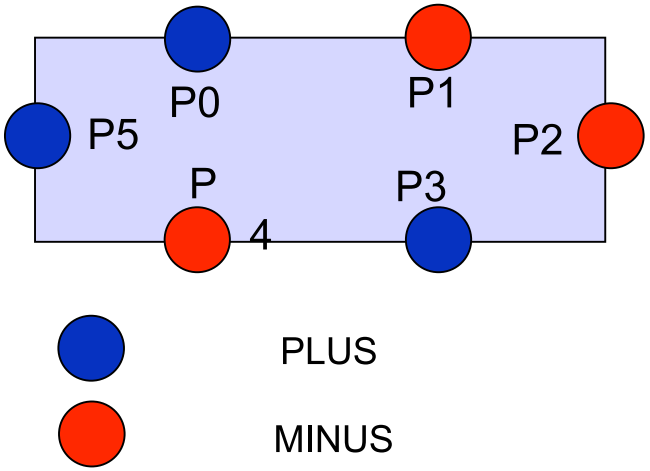
- Job is created on the system in QUEUED status
  - Arguments can be added at this point, if they weren’t added at job creation time
- Startjob command or setJobState API call moves job to STARTING
  - This status value tells to ciodb to start the job, provided the block is not already busy running another job
- Ciodb contacts the IO nodes with the job and user information, starts the job and status moves to RUNNING
  - Ciodb polls the job for stderr, stdout, and completion
- When ciodb is told by ciod that the job has ended, job status goes to TERMINATED
  - Job record is moved from active job table to job history table
- If a user kills the job prior to completion, there is an intermediate status of DYING
  - This notifies ciodb to kill the job, and then set the status to TERMINATED

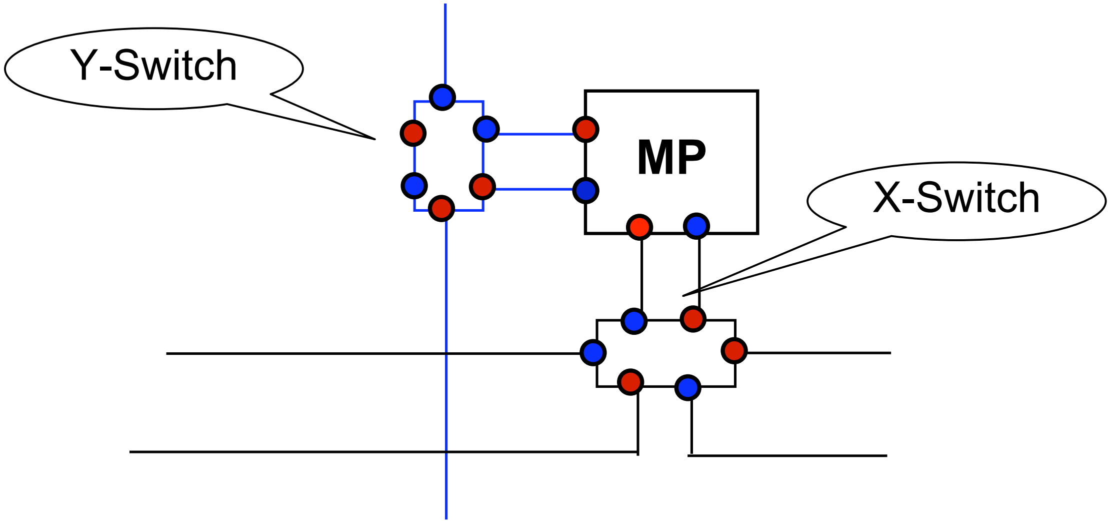
Partitioning of BG/L

- The control system manages all aspects of hardware partitioning, using cables and link chip programming
- The control system handles link programming by making “switch settings” during the boot process
- Allows many simultaneous jobs to be running
- Each partition is isolated
- Partitions can be running with different kernels

The Switches

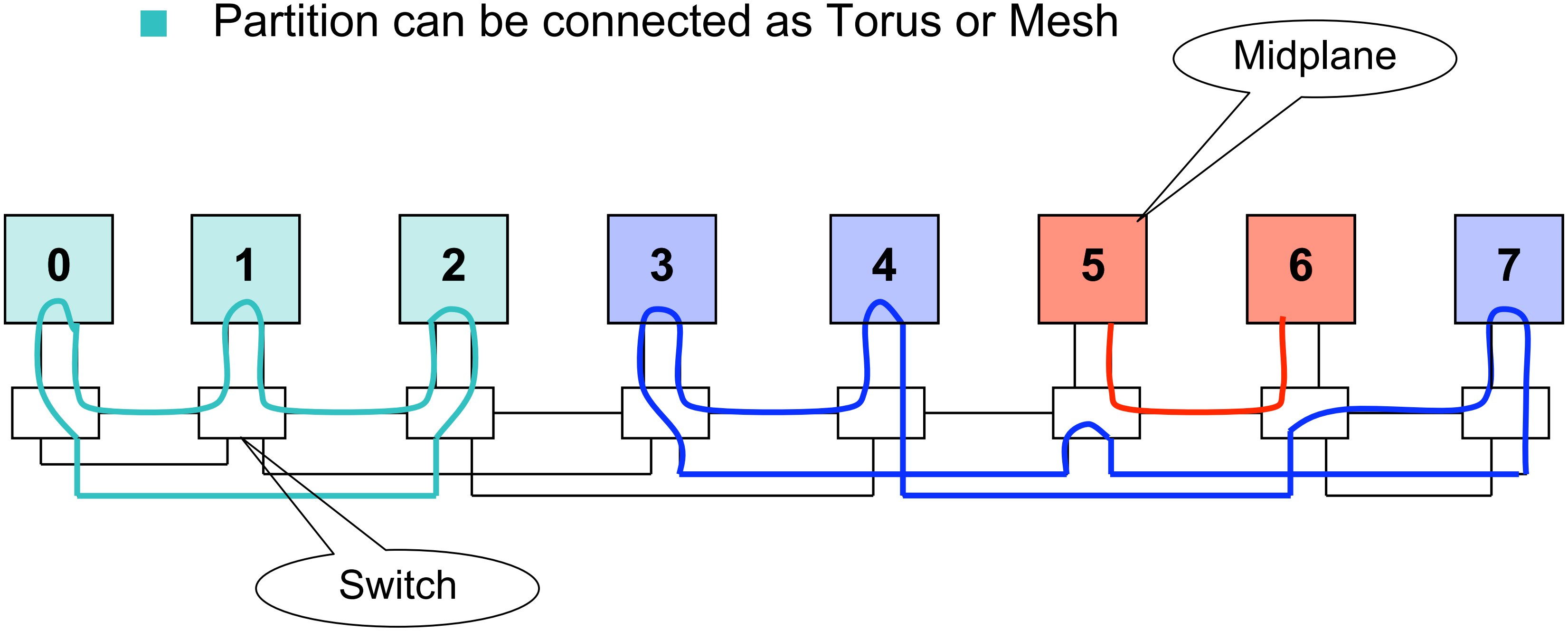
- Each Midplane is connected to three switches
  - One switch on each dimension (X/Y/Z)
- Each switch has six ports (P0..P5)
  - Two ports connect to the midplane (P0,P1)
  - Other four connect to other switches (P2..P5)
- No direct connection between switches on different dimensions



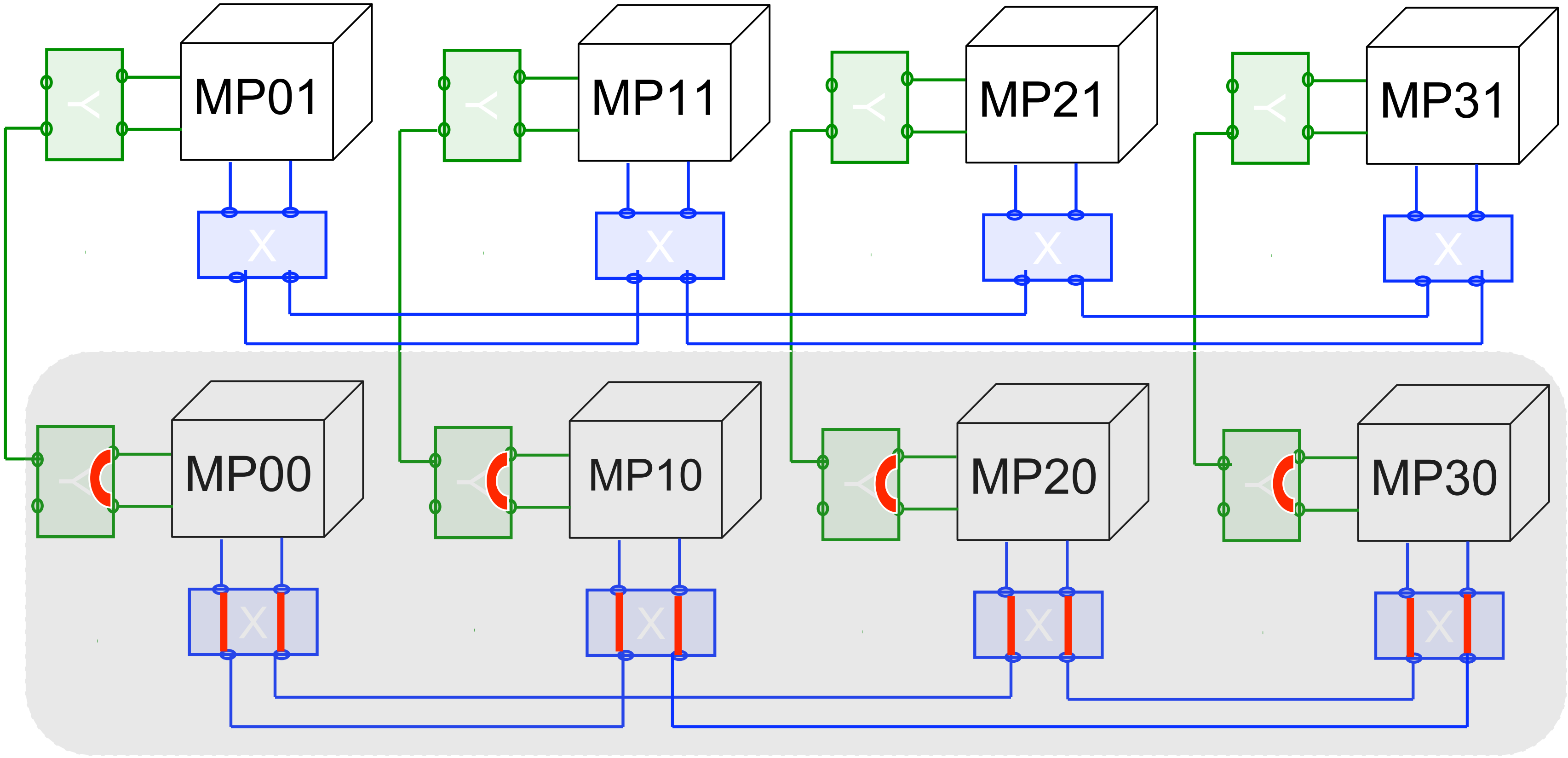


Partition Allocation on Multiple Midplanes – 1D Example

- Partitions are allocated in an isolated manner
  - No Congestion
  - Enhanced Security
- Partition can be connected as Torus or Mesh



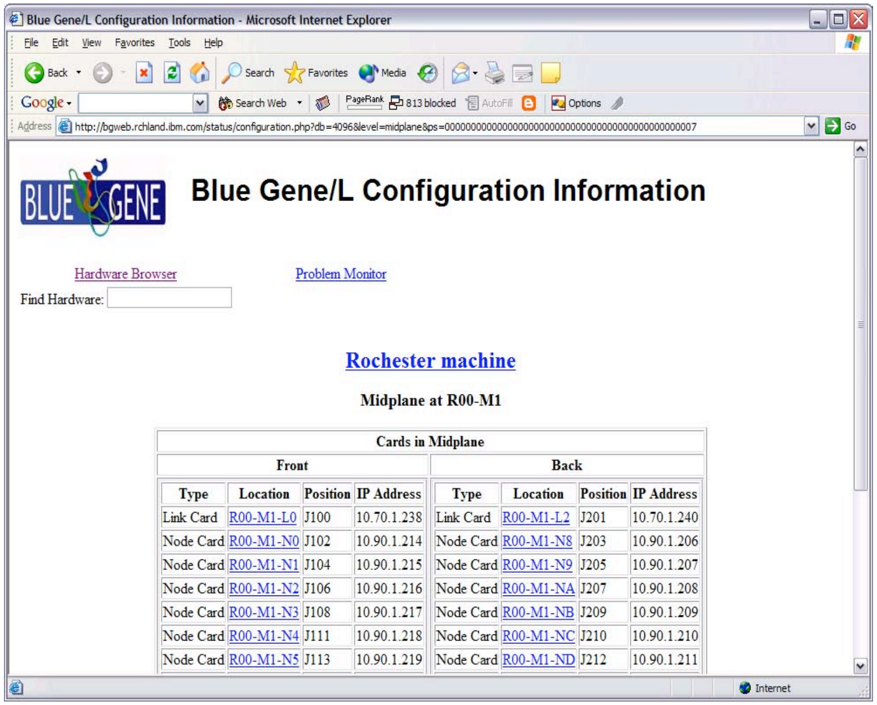
Example – A 2D Machine



Web Interface to DB

- A front-end that runs via browser to view DB2 data.
- Supports the viewing of RAS data, configuration data, diagnostics data, and operational data.
- Can be used to see how the hardware fits together
- Can be used to find trouble areas, hardware anomalies
- Eliminates the need to have SQL expertise to view basic BlueGene information from the database.

Web Page Screen Shots


Blue Gene/L Configuration Information - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Google Search Web PageRank 813 blocked AutoFill Options
Address
Go
Internet

http://bgweb.rchland.ibm.com/status/configuration.php?db=4096&level=midplane&ps=000000000000000000000000000000000000000000007

Blue Gene/L Configuration Information

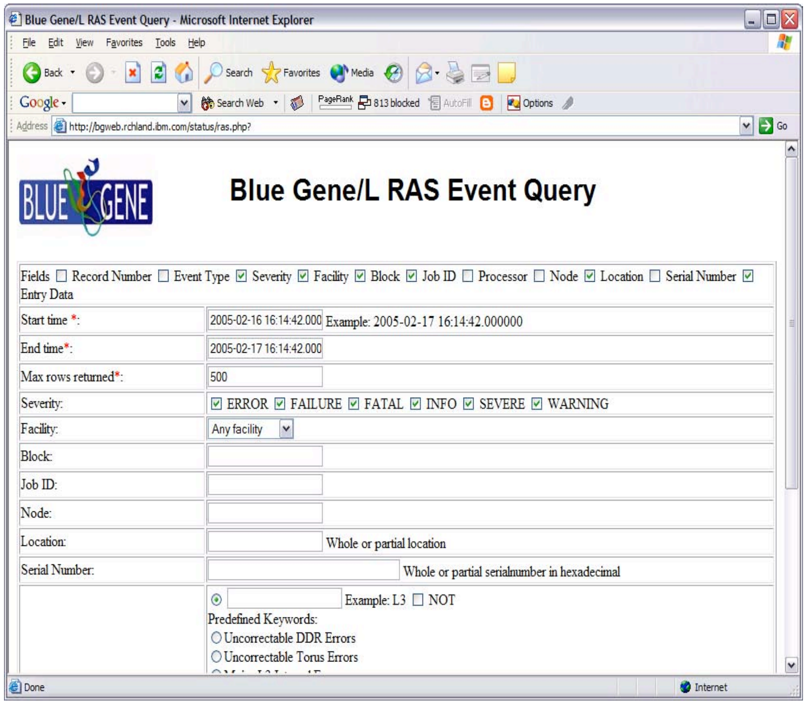
Hardware Browser   Problem Monitor

Find Hardware:

Rochester machine

Midplane at R00-M1

Cards in Midplane							
Front				Back			
Type	Location	Position	IP Address	Type	Location	Position	IP Address
Link Card	R00-M1-L0	J100	10.70.1.238	Link Card	R00-M1-L2	J201	10.70.1.240
Node Card	R00-M1-N0	J102	10.90.1.214	Node Card	R00-M1-N8	J203	10.90.1.206
Node Card	R00-M1-N1	J104	10.90.1.215	Node Card	R00-M1-N9	J205	10.90.1.207
Node Card	R00-M1-N2	J106	10.90.1.216	Node Card	R00-M1-NA	J207	10.90.1.208
Node Card	R00-M1-N3	J108	10.90.1.217	Node Card	R00-M1-NB	J209	10.90.1.209
Node Card	R00-M1-N4	J111	10.90.1.218	Node Card	R00-M1-NC	J210	10.90.1.210
Node Card	R00-M1-N5	J113	10.90.1.219	Node Card	R00-M1-ND	J212	10.90.1.211


Blue Gene/L RAS Event Query - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Google Search Web PageRank 813 blocked AutoFill Options
Address
Go
Done
Internet

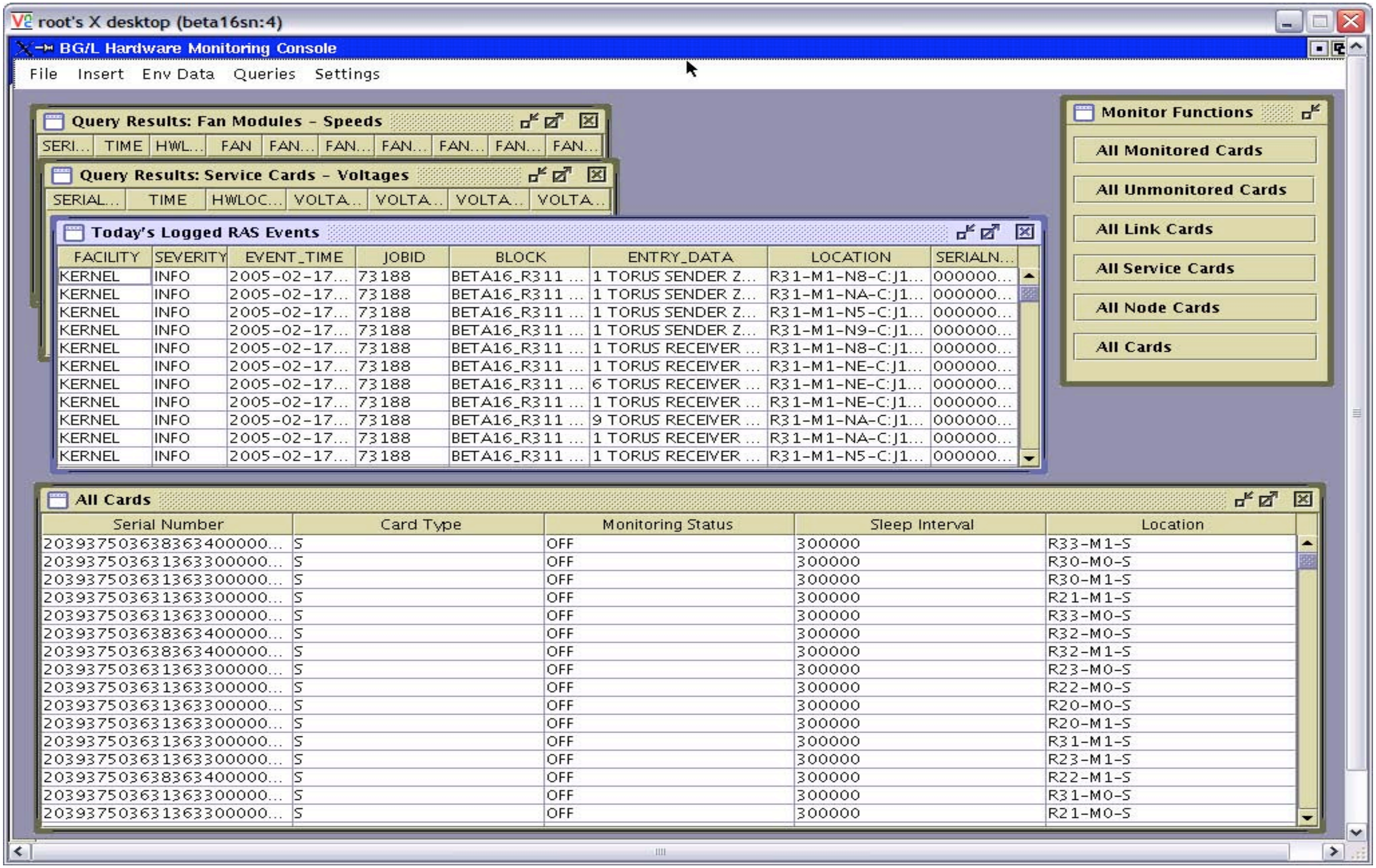
http://bgweb.rchland.ibm.com/status/ras.php?

Blue Gene/L RAS Event Query

Fields □ Record Number □ Event Type ☑ Severity ☑ Facility ☑ Block ☑ Job ID □ Processor □ Node ☑ Location □ Serial Number ☑ Entry Data

Start time *:	2005-02-16 16:14:42.000 Example: 2005-02-17 16:14:42.000000
End time*:	2005-02-17 16:14:42.000
Max rows returned*:	500
Severity:	☑ ERROR ☑ FAILURE ☑ FATAL ☑ INFO ☑ SEVERE ☑ WARNING
Facility:	Any facility
Block:	
Job ID:	
Node:	
Location:	Whole or partial location
Serial Number:	Whole or partial serialnumber in hexadecimal
	● Example: L3 □ NOT Predefined Keywords: ○ Uncorrectable DDR Errors ○ Uncorrectable Torus Errors

Hardware Monitor Screen Shot



Query Results: Fan Modules - Speeds

SERI...	TIME	HWL...	FAN	FAN...	FAN...	FAN...	FAN...	FAN...	FAN...

Query Results: Service Cards - Voltages

SERIAL...	TIME	HWLOC...	VOLTA...	VOLTA...	VOLTA...	VOLTA...

Today's Logged RAS Events

FACILITY	SEVERITY	EVENT_TIME	JOBID	BLOCK	ENTRY_DATA	LOCATION	SERIALN...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS SENDER Z...	R31-M1-N8-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS SENDER Z...	R31-M1-NA-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS SENDER Z...	R31-M1-N5-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS SENDER Z...	R31-M1-N9-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS RECEIVER ...	R31-M1-N8-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS RECEIVER ...	R31-M1-NE-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	6 TORUS RECEIVER ...	R31-M1-NE-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS RECEIVER ...	R31-M1-NE-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	9 TORUS RECEIVER ...	R31-M1-NA-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS RECEIVER ...	R31-M1-NA-C:J1...	000000...
KERNEL	INFO	2005-02-17...	73188	BETA16_R311 ...	1 TORUS RECEIVER ...	R31-M1-N5-C:J1...	000000...

All Cards

Serial Number	Card Type	Monitoring Status	Sleep Interval	Location
203937503638363400000...	S	OFF	300000	R33-M1-S
203937503631363300000...	S	OFF	300000	R30-M0-S
203937503631363300000...	S	OFF	300000	R30-M1-S
203937503631363300000...	S	OFF	300000	R21-M1-S
203937503631363300000...	S	OFF	300000	R33-M0-S
203937503638363400000...	S	OFF	300000	R32-M0-S
203937503638363400000...	S	OFF	300000	R32-M1-S
203937503631363300000...	S	OFF	300000	R23-M0-S
203937503631363300000...	S	OFF	300000	R22-M0-S
203937503631363300000...	S	OFF	300000	R20-M0-S
203937503631363300000...	S	OFF	300000	R20-M1-S
203937503631363300000...	S	OFF	300000	R31-M1-S
203937503631363300000...	S	OFF	300000	R23-M1-S
203937503638363400000...	S	OFF	300000	R22-M1-S
203937503631363300000...	S	OFF	300000	R31-M0-S
203937503631363300000...	S	OFF	300000	R21-M0-S

BG/L Control System Summary

- Control system software runs on the service node
- Control system components
  - MMCS
  - Discovery
  - Ciodb
  - Hardware Monitor
  - Proxy
  - RAS
- Control system handles:
  - Partitions
  - Jobs
  - Discovering and Monitoring Hardware
- DB2 is the central repository of all control system information